DOCUMENT RESUME

ED 079 191 SO 005 921

AUTHOR	Postlethwaite, T. Neville
TITLE	Some General Findings from the IEA Surveys in Science, Reading Comprehension, and Literature.
PUB DATE	73
NOTE	5p.; Paper presented at the American Educational Research Association meeting in New Orleans, February 27, 1973
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Cognitive Measurement; Comparative Analysis; *Comparative Education; *Educational Research; Family School Relationship; Interest Tests; Literature; *National Surveys; Reading Tests; Research Methodology; Response Mode; Science Tests; Speeches; Student Teacher Relationship; Testing
IDENTIFIERS	*American Educational Research Association

ABSTRACT

In this paper some general findings from educational surveys in science, reading comprehension, and literature are reported. A brief description of the research methodology is given. Science was tested in 19 countries, reading in 15 countries and literature in 10 countries. The target populations in each country are defined. Outcome measures in science and reading were primarily of a cognitive nature, whereas in literature about half were cognitive and the other half were modes of response to literature and interest measures. Highlights of the data analysis and results of the project are noted in this conference speech presented prior to publication of the detailed IEA Six-subject Survey. (SHM)

Postlethwaite, T. Neville

ED 079191

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Some general findings from the IEA surveys in Science, Reading Comprehension and Literature

A contribution to a Symposium on the IEA Six-subject Survey at the AERA Meeting in New Orleans on 26 February, 1973

INTRODUCTION

The survey, some of the results of which I am presenting today, was conceived in 1964. The instrument construction took place during the period 1966-69; a trial run took place in 1969 and the final testing during 1970.

In Science and Reading, the outcome measures were primarily of a cognitive nature, whereas in Literature about half were cognitive and the other half consisted of modes of response to literature and interest measures. The independent measures consisted of some 500 variables concerning the student himself and his home background, the instruction he received in school, information on the science teachers and the mother tongue teachers, as well as information on the school principal and the school organisation and financing.

Science was tested in 19 countries, Reading in 15 countries and Literature in 10 countries. The countries represented were: Australia, the Flemish part of Belgium, the French part of Belgium, Chile, England, Federal Republic of Germany, Finald, France, Hungary, India, Iran, Italy, Japan, The Netherlands, New Zealand, Scotland, Sweden, Thailand and the United States.

SP005 921

The target populations identified in each country were : (a) 10-year olds primarily because this was the last point in most countries where students had a class teacher and before they moved to subject-specialised teachers (designated Population I); (b) 14-year olds in full-time schooling, because this was the last point in most systems where 100% of an age group was still in school and before major drop-out began (designated Population II); and, finally, (c) all students in the pre-university year, i.e. the terminal grade of secondary school (designated Population IV).

Probability proportional to size (PPS) samples were drawn from each of the populations. The sampling design was a complex one and was either 2-stage (i.e. school and then student within school), or 3-stage (area, schools within area, students within schools). The average size of sample for any one subject for any one population for any one country was approximately 100 schools and 3,000 pupils. This gives some idea of the magnitude of the study.

AERA New Orleans '73

FILMED FROM BEST AVAILABLE COPY

In general, the sampling designs were well-executed - in some cases excellently so - and the testing exercise went off well.

Standardised manuals of instruction had been prepared for :

(a) national research centres
(b) school co-ordinators, and
(c) test administrators.

The national centres had been given standard forms for returning data to the IEA headquarters: either on MRC cards, on punched cards or in card image form on magnetic tape. Unfortunately, not all national centres stuck to the precise instructions, with the result that when the data were received, a good deal of extra programming had to be undertaken to cope with some of the national idiosyncrasies and to bring them back into the standard form. Indeed, data processing proved to be the greatest headache of all the steps in the whole project. At one point the data processing was nine months behind schedule. On the other hand, if one considers that the final analyses for Science, Reading Comprehension and Literature were run in December, 1972, for data which were received in the period October 1970 to April 1971, and that the magnitude of the study was enormous, then the data processing compares very favourably with that of other projects.

Before presenting some of the general results, it ought perhaps to be re-emphasised that the project was of a co-operative nature, that the test construction involved collaborative work from teams of subject-matter specialists in each of the countries concerned and an international committee, and that agreement was reached on the general appropriateness of the tests to all the nations concerned. This is not to say that the tests were one hundred per cent appropriate to the curriculum of any one country.

The reports have now been written and have gone to press. They will be published in May, 1973, by Almqvist and Wiksell of Stockholm, conjointly for the United States with John Wiley & Sons, Inc..

SOME RESULTS

Certain analyses have been undertaken, particularly in Science, to indicate the relative strengths and weaknesses in the performance of students in Science in the various countries, and an attempt has been made to relate these performance results to centralised curriculum plans where they exist. The analyses are detailed and I will not give results here, but only mention that they have been undertaken and reported.

There is a considerable range in the country mean scores and, in general, the developing countries score one to one and a half standard deviations below the mean for developed countries.

At this juncture it should perhaps be mentioned that a reading speed test was given as part of the Reading Comprehension test battery. The Reading Speed Test consisted of 40 items, each of approximately the following difficulty and format:

"Peter has a little dog. The dog is black with a white spot on his back and one white leg. The colour of Peter's dog is mostly

black brown grey"

Each student had also received a practice test of the same length immediately before undertaking the actual speed test. Reading speed was to be measured by the item reached.

It so happened that on the first page of the test, in every country, there were the first 9 items. These items were scored for correctness and the number of errors was recorded. In the developing countries there was a relatively high correlation between the number of errors on the first 9 items and the item reached, which suggested a haphazard form of answering.

In the European countries, a typical error rate on these items was about 10% for 10-year olds and 4% for 14-year olds. In the developing countries, the error rate was ranging between 26-52% for 10-year olds and 16-33% for 14-year olds.

If a substantial proportion of the students in a school system have real difficulty in reading these materials, one must question whether any more than a minimal level of literacy has been achieved in that school system. Furthermore, it casts doubt upon the testing exercise in certain developing countries. Could they even read the Science tests or the background questionnaires? To what extent are some of the odd results in a multivariate analysis attributable to non-comprehension of the materials, or do they in fact reflect the reality in those school systems? Is there a case for a detailed study in one or two developing countries on how to collect data reliably at a relatively low cost?

THE HOME vs. THE SCHOOL

In undertaking the multivariate analysis, a multiple regression technique was used. The order of entry of variables was discussed over 4-5 years by eminent statisticians and educators. Agreement proved difficult to attain but in the end the first block of variables to be entered in the regression consisted of long-term home background variables which were likely to have remained relatively unchanged since the birth of the child.

The second block of variables consisted of the type of school or type of programme in which the child was enrolled, these being thought to be reasonable surrogates for the previous educational experiences of the child. The third group of variables were the teacher and school variables. The fourth group were variables which were considered to be associated or kindred variables, but not causal variables. As can be seen, this is a temporal model whereby in education what happens yesterday can affect what happens today, but what happens today cannot affect what happened yesterday. Within each of the blocks, a stepwise multiple regression was used without a pre-determined order on the variables in the block.

The findings are somewhat different from those given in the Coleman "Equality of Educational Opportunity" report. However, it must be borne in mind that the dependent variables in Coleman's report were mostly Reading and a little bit of Arithmetic.

The home (whether on a between student or a between school basis) always accounts for more variance than does the school. This is true for all populations in Reading. Indeed, except at the pre-university grade level, the home is accounting for about 3 times as much variance as the school on the between student basis and 3 to 8 times as much on the between school basis. However, in Science, the higher one moves up the school system, the more important the school variables become and indeed, at the pre-university level, the school is, on average, more important than the home and, in some countries, is accounting for over half the total variance accounted for (in one case it is 41%). It must be remembered, however, that in most countries there has been considerable drop out before the pre-university grade and the variance of home backgrounds is restricted.

On a between school basis, the school is accounting for as much as 45% of the variance. This particular case is a developing country where the total between school variance accounts for 70% of the total between student variance.

In Literature, too, the school is relatively more important at Population IV level than it is at Population II. Could this be that, for more school-oriented subjects, the school is able to account for more variance?

However, it should be noted that, in the United States, the pattern is much the same as Coleman described for each of the subjects tested at each level.

SCHOOL AND TEACHER VARIABLES

If one examines the minimum contribution of school and teacher variables in Block 3 (i.e. entered last at the end of Blocks 1, 2 and 3), the following general picture emerges:

The actual curriculum as measured by the IEA variable "opportunity to learn" is important, i.e. if you allow the students to learn more they, in general, will, and if you don't, they can't.

The number of hours (and years) of instruction received is powerful, as is the number of hours of homework per week.

In Science, the sex of teacher is important and, at Population IV level, the pre-service teacher training.

There are approximately another 20 variables in Science coming through at the final regression stage, but in a haphazard manner. In Reading and Literature, only about another 6 to 10 variables are emerging, but again in an inconsistent way.

However, with 500 independent variables and in the between schools analysis only 100 observations, a great deal of discarding and compositing of variables had to be undertaken. It is clear that it is impossible (or more or less impossible) to speak clearly about the importance of any one specific variable surviving the regression analysis, since those surviving are obviously, at the survival stage, acting to some extent as surrogates for those discarded. Regression analysis is, after all, a rather blunt instrument when attempting to make recommendations about specific variables, given the constraints of the degrees of freedom.

I hope that these tit-bits of findings have proved to be tantalising enough for you to delve further into the detailed analyses and results to be published shortly.

T. Neville Postlethwaite
International Institute for
Educational Planning, Paris